Vectors: An Introduction for Physicists

J.W. Krol

September 12, 2016

1 Numbers

In science we study *quantities*. A quantity has a *magnitude*. For two magnitudes A and B we have either:

- A is smaller than B
- A is greater than B
- A is equal to B

A specific magnitude U is chosen as a standard unit for a physical quantity. A measurement is an operational process to determine the ratio of a specific magnitude M versus the unit magnitude U. This ratio is a dimensionless *number* or *scalar*. It is by measurement that a relation is established between the real world of magnitudes and the abstract world of numbers.

In human history the first form of measurement was simply counting whole things. The most 'natural' way to count is by making a mark for every item counted. If you have three sheep, you write III. If you have two sheep, you write II. If you have no sheep at all, what do you write? Nothing at all! This last observation makes clear that the number 0 we are all familiar with is not so naturally at all.

Although the Greek were well aware of magnitudes and their relations in geometry, they did not assign numbers to them. The Greeks worked with magnitudes of various kinds, including line segments, areas and volumes. Greek mathematicians did not use numbers to measure geometric magnitudes. Numbers were reserved only for counting objects. They did not refer to the length of a segment, but instead compared line segments as ratio's of magnitudes The Greeks constructed new magnitudes out of smaller magnitudes by multiples:

$kM = M + (k-1)M$ or from lager magnitudes by taking parts: $\frac{1}{k}M$ with $k = 1, 2, 3, ...$ These operations of addition and dividing exists without numbers.

It was not until the Renaissance that the role of magnitudes of line segments was taken over by numbers. The decimal number system was invented which finally allowed measuring any magnitude with numbers and defined the basic rules of arithmetic. The decimal number system consists of ten basic symbols $0 \cdots 9$ and from these ten symbols we define whole and fractional numbers. We picked arbitrarily 10 symbols for our number system, but any whole number greater or equal than 2 is fine to set up a number system. Each number is represented by the symbols, the whole parts left to the dot and the fractional parts right to the dot:

$$a_n \cdots a_2 a_1 . a_{-1} \cdots a_{-m} = a_n b^{n-1} + \cdots + a_2 b^1 + a_1 b^0 + a_{-1} b^{-1} + \cdots + a_{-m} b^{-m}$$

where $a_i \in \{0, \cdots, b\}$.

Vieta and Descartes in the 17th century used the decimal number system and contributed further to mathematical insight to generalize arithmetic into algebra. In algebra numbers are represented by letters which are abstract symbols not instantiated to any specific number. This enabled then to extend the numbers from integers to algebraic numbers. Algebraic numbers are numbers that are a solution of an algebraic equation. Finally we could $\sqrt{2}$ define as the number of the solution of the algebraic equation $x^2 = 2$.

Descartes also merged Euclidean geometry with algebra to form analytic geometry. Analytic geometry defines a one-to-one correspondence with directed line segments and numbers. We can therefore solve geometric problems with calculations based on numbers.

The equivalence between line segments and numbers shows also the elementary algebraic properties that must be obeyed by numbers. Let a, b and c be length of three line segments then it is obvious from the geometry of addition and multiplication of line segments that for *numbers* a, b, c the following laws of algebra must hold:

$a + b = b + a$ commutative law of addition

$a + (b + c) = (a + b) + c$ associative law of addition

$ab = ba$ commutative law of multiplication

$a(bc) = (ab)c$ associative law of multiplication

$a(a + b) = ab + ac$ distributive law of multiplication

Once the connection between line segments and numbers had been established and the decimal number system was available the next step in the development of science was taken with the definition of a universal method of measuring any number P on a line.

The measurement of a point P on a line requires to create a numbered line, a *ruler*, with the following conventions:

- a point O, the *origin*, on the line is chosen and assigned the coordinate 0
- a point I, the *unit*, on the line is chosen and assigned the coordinate 1
- one side of the line relative to O is labeled $-$ and the other side $+$
- each point P on the line is given a coordinate $x = x(P)$ determined by the distance of P from O, $|OP|$, measured in unit segments OI and assigned a + or a - based on the side were P is located.

The measurement of P is the determination of the coordinate of P and proceeds as follows:

- lay off the maximum number $a_1 a_2 ... a_m$ of whole unit segments OI in the direction towards P until a point Q. If Q=P stop otherwise go to the next step.
- from Q divide the consecutive unit segment OI into 10 equal parts and move forward b_1 parts until Q^1, if $Q^1 = P$ stop otherwise repeat the foregoing step (infinitely many if needed), now dividing the consecutive unit segment OI into 10^{k+1} equal parts (where k is the number of the previous step) and move forward b_k parts, until $Q^{k+1} = P$.

The point P on the line constructed by the above protocol is represented by the decimal number

$$\pm a_1 a_2 ... a_m . b_1 b_2 .... b_n ....$$

where $\pm$ designates its orientation, a and b are digits from zero to nine representing the integral and decimal part of the number, and m and n are natural numbers indicating the position of the decimal.

Whenever the number has infinitely many non repeating decimals the number is *irrational.* All other numbers are *rational numbers* and can be expressed as a ratio of two integer numbers $\pm\frac{p}{q}$. The aggregate of all numbers, rational and irrational, is called the *arithmetical continuum* of *real numbers* denoted with $\mathbb{R}$. The arithmetical continuum is equivalent to a straight line. The aggregate of the points on a straight line constitute the *linear continuum* and is a convenient image of the arithmetical continuum.

Some rational numbers have like irrational numbers an infinite expansion of decimals (however with some repetition pattern). For both of these numbers the above procedure never stops and runs forever. It caused mathematicians to introduce the concepts of *infinity* and *limit* into their analysis. For an applied science numerical calculations will always be carried out in a finite number of steps. It give rise to the domain of *error analysis* were one studies the impact of rounding on the outcomes of calculations.

In arithmetic and algebra we apply rules to numbers without reference anymore to the geometry of the line. However any physical interpretation of a magnitude of a quantity requires the geometry of the construction of a point on a line. But each point has a one-to-one correspondence with a real number. Whenever we assure that the geometrical rules of construction are aligned with the algebraic rules of operation we can safely use them interchangeably.

The construction and the analysis of the properties of $\mathbb{R}$ is what has been studied extensively by mathematicians starting in the seventeenth century by Newton and Leibniz and continued till about 1900. It is called *Analysis* in the curriculum of university students.

Here our focus lies on a multi-dimensional structure, $\mathbb{R}^n$. The study of $\mathbb{R}^n$ is normally called *vector analysis*. However it is only a generalisation of the analysis of $\mathbb{R}$ into higher dimensions.

2 Geometry

We draw a *right triangle*, consisting of a diagonal, a horizontal line segment X with length x and a vertical line segment Y with length y. X and Y perpendicular to each other. A triangle expresses a relationship between *distance* and *angles*. It is an important geometrical object as it specifies the relation between length and rotation.

We can define operational procedures to measure a distance along a straight line with a *metre stick* and measuring angle with a *goniometer*. However mathematics allows us to calculate distances and angles based on points only. To do that we need to further develop our knowledge about these relations.

Let us first define more precisely an angle. To define an angle we follow the convention that the angle is the acute angle [1]. If the turn is counterclockwise the angle is positive and else negative. An angle between two straight lines, l and m which intersects at O, is defined as the ratio of the *arc length*, s, between a point A on l and B on m each at a distance r from O, and the circumference of a circle of radius r:

$$\theta = \frac{s}{r}$$

A radian is an angle subtended by an arc whose length is equal to the radius. Radian is a dimensionless quantity as it is the ratio of two lengths. However we use the symbol rad to denote that the angle is measured in radians to avoid confusion with degrees.

[1] angle less than or equal $90°$

For a full circle we have:

$$\theta = \frac{2\pi r}{r} = 2\pi \text{ rad}$$

And therefore the following conversion equation:

$$1 \text{ rev } = 360^\circ = 2\pi \text{ rad}$$

The relations between distance and angles depend on the specific geometry of the space. The geometry most common is the **Euclidean geometry**, also called flat geometry, which is based on 5 postulates:

- A straight line segment can be drawn joining any two points.
- Any straight line segment can be extended indefinitely in a straight line.
- Given any straight lines segment, a circle can be drawn having the segment as radius and one endpoint as center.
- All Right Angles are congruent.
- If two lines are drawn which intersect a third in such a way that the sum of the inner angles on one side is less than two Right Angles, then the two lines inevitably must intersect each other on that side if extended far enough.

The last postulate is called the Parallel Postulate. In history mathematicians tried to deduce this postulate from the 4 others, but failed. It turned out that an equivalent formulation of this postulate is:

Through any given point can be drawn exactly one straight line parallel to a given line.

This has lead to the insight that when one reformulates this postulate a new *non-Euclidean geometry* results:

Spherical Geometry: *Through any given point can be drawn no straight line parallel to a given line.*

Hyperbolic Geometry: *Through any given point can be drawn more than one straight line parallel to a given line.*

Here we will assume an Euclidean geometry of space. The important point is that there is not one geometry and once we have chosen a geometry to work with, a method is determined to assign numbers to each point of the space, which are called coordinates, and a method to calculate a distance between points. The form of the metric depends on the coordinate system chosen and the geometry of the space. In our case we have a rectangular coordinate system and an Euclidean geometry.

Pythagoras Theorem Based on the definitions and axioms of Euclidean geometry one can proof a statement about the length of the diagonal of a rectangular triangle. It is known as *Pythagorean Theorem*. This theorem states that the length of the diagonal of a right triangle squared is the sum of its legs squared:

$$c^2 = a^2 + b^2$$

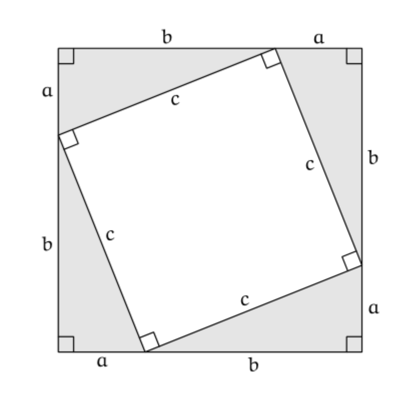

Figure 1: Proof of Pythagoras Theorem

Proof 1 (Pythagoras)**.**

$$\begin{aligned}(a+b)^2 &= c^2 + 4\frac{ab}{2} \\ a^2 + 2ab + b^2 &= c^2 + 2ab \\ c^2 &= a^2 + b^2\end{aligned}$$

Cosine rule The relation of the length of the sides of a triangle ABC can be generalized to any triangle with the cosine rule:

$$a^2 = b^2 + c^2 - 2bc\cos A$$

where a, b, c are the lengths of the sides opposite to the vertices A, B, C respectively and $\cos A$ the angle at vertex A.

Proof 2. *Given a triangle ABC with an acute angle A, see figure 2. We have $x = b\cos A$. We deduce*

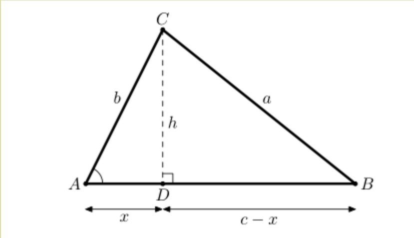

Figure 2: Proof cosine rule acute angle

$$
\begin{aligned}
(c-x)^2+h^2 &= a^2\\
a^2 &= c^2-2cx+(x^2+h^2)\\
a^2 &= c^2-2bc\cos A+b^2\\
a^2 &= b^2+c^2-2bc\cos A
\end{aligned}
$$

Given a triangle ABC with an obtuse angle A, see figure 5. We have $x=-b\cos A$. *This follows from* $\cos\pi - A = -\cos A$.

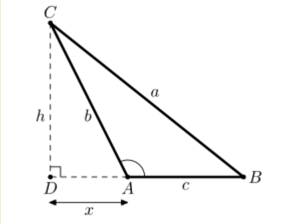

Figure 3: Proof cosine rule obtuse angle

$$
\begin{aligned}
(c+x)^2+h^2 &= a^2\\
a^2 &= c^2+2cx+(x^2+h^2)\\
a^2 &= c^2-2bc\cos A+b^2\\
a^2 &= b^2+c^2-2bc\cos A
\end{aligned}
$$

3 Vectors

3.1 Vectors in a plane

On a line we have one quantity to measure: length. In a plane there are two quantities: length and angle. In a plane any three points, non-collinear, define two intersecting directed line segments. If we define the point of intersection, O, as the initial point of the two directed line segments, then we can measure the lengths of both directed line segments as well as the angle between them. If we choose one of these directed line segments as a basis, then for any point P of space, a directed line segment OP is defined. This directed line segment is uniquely represented by the point O and two numbers, length and angle.

A directed line segment can be translated to any location in the plane, and as a translation preserves the length and angle, each of these translated directed line segments is then represented by the same two magnitudes of length and angle with the defined basis. The only difference between any of these directed line segments is the location in space.

We define a *vector* in a plane as a mathematical object, represented by two numbers which represent two quantities length and angle. A vector then represents all directed line segments in space that coincident with each other after a translation or stated otherwise a vector is the set of all parallel directed line segments with the same length and angle. A vector is another kind of number, it has no location in space, only points have.

We denote a vector symbolically as follows: $\vec{x}$ or $\boldsymbol{x}$ (bold letter), the magnitude of the two quantities are called *polar coordinates* of the vector $\boldsymbol{x} = (r, \theta)$. The coordinates are dependent on basis we have chosen, however the vector itself is independent of this choice. The length of a vector denoted as $\|\mathbf{x}\|$.

If $\boldsymbol{x}$ is a vector and P a point in space then $\boldsymbol{x}$ translates P to the point P'. A vector is like a step in space from P to P'. This is in analogy with a number which translates the point O to a point P on the line. A vector translates a point in space.

3.2 Vectors in n dimensional space

The concept of a vector is not limited to a plane. In 3 dimensional space we choose again three points and measure two numbers, length and angle, in the way as described before. However these two numbers do not uniquely represent a vector anymore, as we can rotate each terminal point of such a vector around the third dimension between $0° - 360°$. Any of these rotated vectors have the same length and angle with respect to the original basis, but have a different angle with respect to the third dimension and therefore have a different direction

in 3 dimensional space. To uniquely quantify a directed line segment in 3 dimensional space we need also to fix an angle with another basis directed line segment non-collinear with the first. We need 4 points non-coplanar to form the basis vectors. We measure for each directed line segment the angles with these two basis vectors. A vector in 3 dimensional space is thus represented by three numbers: a length, and two angles with two basis vectors $\boldsymbol{x} = (r, \theta, \phi)$. These coordinates are called *spherical coordinates*.

Of course the concept can be extended in the same way to any n dimensional space. Each additional dimension introduces an additional axis of rotation and we need a length and $n-1$ angles to represent a vector in n dimensional space.

3.3 Addition and Scalar multiplication

A vector represent geometrically a translation in space. The basic operations on a vector are the addition of vectors and the multiplication of its length.

- There is a zero vector $\boldsymbol{o}$ which translates P to P.
- Adding $\boldsymbol{x} + \boldsymbol{y}$ results in a vector $\boldsymbol{z}$ which is the combined translation of $\boldsymbol{x}$ and $\boldsymbol{y}$. This property is called the *triangle law* of addition. Adding $\boldsymbol{x} + \boldsymbol{y}$ is equivalent to $\boldsymbol{y} + \boldsymbol{x}$, the operation is **commutative**. Also $\boldsymbol{x} + (\boldsymbol{y} + \boldsymbol{z})$ is equivalent to $(\boldsymbol{x} + \boldsymbol{y}) + \boldsymbol{z}$, the operation is **associative**. This property is called the *parallelogram law* of addition.
- Multiplying a vector $\boldsymbol{x}$ with a number λ results in a vector $\boldsymbol{y}$ whose magnitude equals: $\lambda \left\| \mathbf{x} \right\|$. $\lambda(\boldsymbol{x} + \boldsymbol{y}) = \lambda \boldsymbol{x} + \lambda \boldsymbol{y}$, the operation is distributive. If we multiply a vector $\boldsymbol{x}$ by the inverse of its length $1/\left\| \mathbf{x} \right\|$ a unit vector results: $\hat{\boldsymbol{e}}_x$, this unit vector is parallel to $\boldsymbol{x}$ and has length 1. Multiplying a vector with -1 reverse the direction of the vector, an 180° rotation of the angle of the vector. Substraction of a vector $\boldsymbol{x}$ from another vector $\boldsymbol{y}$ is the addition of the inverse of $\boldsymbol{x}$: $\boldsymbol{y} - \boldsymbol{x} = \boldsymbol{y} + (-1.\boldsymbol{x})$.

How can we extend the concept of **multiplication** of two numbers to the multiplication of two vectors ?

3.4 Multiplication

Geometrical the product of two numbers is the area of a rectangle. But an area is just one specific instance of a class of geometrical objects we call **parallelograms**. We associate multiplication of vectors with a parallelogram.

Definition In Euclidean geometry, a parallelogram is a convex [2] quadrilateral (four straight sides) with two pairs of parallel sides. The opposite or facing sides of a parallelogram are of equal length and the opposite angles of a parallelogram are of equal measure.

We define two forms of multiplication: scalar product which produces a number and vector product which produces another vector.

3.4.1 Scalar product

The area of a parallelogram, with sides x and y can be easily deduced as:

$$A = \|\boldsymbol{x}\| \, \|\boldsymbol{y}\| \sin \alpha$$

with α the angle between the sides x and y in radians.

To extend the concept of multiplying numbers to multiplying vectors we can associate the area of the parallelogram formed by two vectors as their product. If $\alpha = \frac{1}{2}\pi$ we have $\sin \alpha = 1$ and the formula reduces to the area of a rectangle similar to multiplication of numbers. So this definition could work. However it turns out that in nature a force (vector quantity) that is orthogonal to the direction of motion (vector quantity) does not contribute to the motion, and when it has the same direction it contributes proportional to motion. If we want the product useful in physics we must substitute α with $\alpha - \frac{1}{2}\pi$ in the formula, then if $\alpha = 0$ we have the maximal value and for $\alpha = \frac{1}{2}\pi$ we have 0. In line with the experience from physics.

Our formula for the *scalar product* of two vectors result:

$$\boldsymbol{x} \cdot \boldsymbol{y} = \|\boldsymbol{x}\|\|\boldsymbol{y}\| \sin\left(\alpha - \tfrac{1}{2}\pi\right) = \|\boldsymbol{x}\|\|\boldsymbol{y}\| \cos \alpha$$

The scalar product returns a number not a vector. The scalar product can be used for any n dimensional space. As $\cos \alpha = \cos -\alpha$ we have $\boldsymbol{x} \cdot \boldsymbol{y} = \boldsymbol{y} \cdot \boldsymbol{x}$.

The component of a vector $\boldsymbol{F}$ in the direction of a vector $\boldsymbol{x}$ is defined as the projection of $\boldsymbol{F}$ onto $\boldsymbol{x}$, which is the scalar product of $\boldsymbol{F}$ and the unit vector of $\boldsymbol{x}$: $\boldsymbol{F} \cdot \hat{\boldsymbol{e}}_x = \|\mathbf{F}\| \cos \alpha$.

From a physics perspective the scalar product multiplies the length of one vector times the length of the component of another vector in the direction of this vector. The result is the rectangular area formed by these two lengths.

[2] A convex polygon is a polygon where no line segment connecting two points on the boundary ever goes outside the polygon or equivalent all angles are less than or equal 180°

3.4.2 Vector product

A product of two vectors that produces a unique vector is called a vector product. Why do we need a vector product in physics ? Rotational motion is caused by a twisting force called *torque*. The torque, τ, on a particle P is proportional to the force $\boldsymbol{F}$ and the vector $\boldsymbol{r}$ between the point of rotation O and P. For this product $\boldsymbol{r} \times \boldsymbol{F}$ we can't use the scalar product defined above, because the torque is maximal if the force is perpendicular to the position vector, in which case the scalar product produces zero. So we could define the vector product based on the original formula of the area of a parallelogram: $\|\mathbf{x}\| \, \|\mathbf{y}\| \sin \alpha$. This area number defines the length of the vector product. Now it remains to define a direction for this product vector. First it often turns out that a vector perpendicular (normal) to a plane is often very useful. Therefore we define the product vector to be normal to the plane formed by $\boldsymbol{F}$ and $\boldsymbol{r}$. Next we must define whether the product vector points outward,$\odot$, or inwards,$\otimes$, from the plane seen from above. The only direction in this case is the direction of rotation of P. P can rotate in either a clock-wise or anti clock-wise direction. A clock-wise rotation occurs whenever the turn from $\boldsymbol{r}$ to $\boldsymbol{F}$ (over the smaller angle between them) is clock-wise, otherwise the rotation is anti-clockwise. As a right-handed screw moves inwards with a clock-wise rotation, and outwards with an anti clockwise rotation, we define the vector product with the *right-hand screw rule* to be inward if the turn from $\boldsymbol{r}$ to $\boldsymbol{F}$ is clock-wise and outward otherwise.

It is good to know that although the torque mathematically behaves like a vector, it is in fact a *pseudo-vector*, invented for our purpose. A pseudo-vector differentiates from a *polar vector*, like position and force, in that its direction is not something we can measure in nature, but is based on a rule of agreement. In physics we use the so called right-hand screw rule. Anytime we refer to this rule we have a pseudo-vector.

We define the cross product of the vectors $\boldsymbol{x}$ and $\boldsymbol{y}$, written $\boldsymbol{x} \times \boldsymbol{y}$, to be a vector perpendicular to the plane defined by $\boldsymbol{x}$ and $\boldsymbol{y}$. The magnitude of this vector is the area of the parallelogram formed by the two vectors. The direction of $\boldsymbol{x} \times \boldsymbol{y}$ is given by the normal vector, $\boldsymbol{n}$. If the turn from $\boldsymbol{x}$ to $\boldsymbol{y}$, along the smaller angles between both vectors, is counterclockwise the orientation of $\boldsymbol{n}$ is outward, otherwise its is inward of the plane formed by $\boldsymbol{x}$ and $\boldsymbol{y}$ (right-hand screw rule):

$$\boldsymbol{z} = \boldsymbol{x} \times \boldsymbol{y} = \|\mathbf{u}\| \, \|\mathbf{v}\| \sin \theta \, \hat{\mathbf{n}}$$

where $0 \leq \theta \leq \pi$ is the angle between $\boldsymbol{x}$ and $\boldsymbol{y}$ and $\hat{\mathbf{n}}$ is the unit vector perpendicular to $\boldsymbol{x}$ and $\boldsymbol{y}$ and in the direction given by the right-hand screw rule. If $\boldsymbol{x}$ and $\boldsymbol{y}$ are parallel then $\boldsymbol{x} \times \boldsymbol{y} = 0$.

3.5 Cartesian Coordinates

To apply the operations given above we must measure distances and angles with a ruler and a goniometer. We will now turn to the arithmetizing of the space by **coordinates** and then we will further analyze the properties of these operations in terms of coordinates, called *vector algebra.*

3.5.1 Plane

We define a *coordinate frame* consisting of two straight lines L_1 and L_2 which intersect at the point O, and define for each line a ruler, x respectively y, each ruler has its zero at the point O. We define on both rulers a unit point, I respectively J. It is convenient, but not necessary to choose these points, so that $OI = OJ$. The tripe (O, I, J) defines the coordinate frame.

On each ruler we need to define a positive and negative direction. We have two rulers, suppose x is horizontal and positive to the right and y is vertical. We have two choices for $y+$, up or down. Mathematicians have agreed upon a convention to choose $y+$ in the *counterclockwise* direction of $x+$. Note that it does not matter in which direction $x+$ points, it is only the relation between $x+$ and $y+$. If we had chosen $x+$ to the left then $y+$ is down. Or if we switch x and y from horizontal to vertical the same rule would apply. If $x+$ is up then $y+$ is to the left and if $x+$ is down then $y+$ is to the right.

If P is any point then it lies on a unique line parallel to L_1 and also on a unique line parallel to L_2. The line parallel to L_1 intersects the line L_2 at at point A, with coordinate $a = x(A)$ and the line parallel to L_2 intersects the line L_1 at the point B with coordinate $b = y(B)$. We then define $[P] = (x(A), y(B)) = (a, b)$ as the coordinates of P, (x, y) is a *homogeneous coordinate system* with *coordinate axes* x and y.

The directed line segment $\overrightarrow{OP}$ is called a *position vector*, because its initial point is aligned with the origin O. We have $\boldsymbol{p} = \overrightarrow{OP}$ and if we define the unit vectors $\hat{\boldsymbol{i}} = \overrightarrow{OI}$ and $\hat{\boldsymbol{j}} = \overrightarrow{OJ}$ we have also $\boldsymbol{p} = x(A)\hat{\boldsymbol{i}} + y(B)\hat{\boldsymbol{j}}$ and we can represent the vector $\boldsymbol{p}$ also by the same coordinates as the point P:

$$\boldsymbol{p} = (x(A), y(B)) = (a, b)$$

The correspondence between points and vectors and pairs of real numbers, as their coordinates, is a one-to-one and onto correspondence (bijective), meaning that every point and vector in the plane corresponds to one pair of numbers, and every pair of numbers corresponds to one point and vector only. It assures us that pairs of real numbers are equivalent to points and vectors in the plane. We adopt the notation $\mathbb{R}^2$ for all pairs of real numbers

in the plane. This bijective relationship is not only valid for two dimensions, but for any number of dimensions. We use $\mathbb{R}^n$ for a tuple of n real numbers.

We will analyze how we can use these numbers to algebraically calculate the magnitude and direction of vectors in a two dimensional space. The magnitude of a geometrical vector is the distance between its initial and end point and the direction is the *angle* between the vector and one of its coordinate axes, we choose here x.

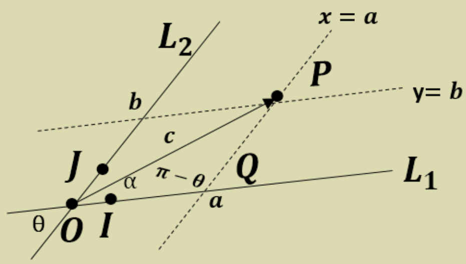

Figure 4: magnitude and direction

To calculate the magnitude of $\boldsymbol{p}$ we choose $\overrightarrow{OP}$ and calculate the length $\overrightarrow{OP}$. We use the angle $\angle PQO = \pi - \theta$ where θ is the angle between the two coordinate axes and apply the *law of cosine* from trigonometry:

$$\|\boldsymbol{p}\| = OP = c = \sqrt{a^2 + b^2 - 2ab\cos(\pi - \theta)} = \sqrt{a^2 + b^2 + 2ab\cos\theta}$$

where θ is the angle between the two coordinate axes, a the x-coordinate and b the y-coordinate of P. Once we have the magnitude we can find the direction from:

$$\cos\alpha = \frac{a}{\|\boldsymbol{p}\|}$$

If we choose $\theta = \frac{1}{2}\pi$ then we have a *rectangular coordinate system* and the equations simplify to the well-known *Pythagoras Theorem* ($c^2 = a^2 + b^2$):

$$\|\boldsymbol{p}\| = \sqrt{a^2 + b^2}$$
$$\cos\alpha = \frac{a}{\sqrt{a^2 + b^2}}$$

when $\theta \neq \frac{1}{2}\pi$ we have an *oblique coordinate system*. Due to its simplicity a *rectangular Cartesian coordinate system* is used often and called an *Cartesian coordinate system*.

The following relations exist between polar coordinates and Cartesian coordinates:

$$x = r\cos\theta$$
$$y = r\sin\theta$$

3.6 3 Dimensional Space

Consider the case of three dimensions. We add to the two dimensional coordinate system a third line l_3 through O not in the plane formed by l_1 and l_2. We define a ruler z for this line with zero at O and a unit point K. It is convenient, but not necessary to choose these unit points such that $OI = OJ = OK$. The quadruple (O, I, J, K) defines the coordinate frame for a three dimensional space.

In the same way as for the two dimensional case we have two choices for the positive direction of z. A right handed three dimensional coordinate system is such that $y+$ is in a counterclockwise direction of $x+$ and $z+$ in a counterclockwise direction of $y+$.

If P is any point then it lies on a unique line parallel to l_3. This line intersects the plane formed by l_1 and l_2, the xy plane, at the point P'. P' has x and y coordinates $a = x(A)$, $b = y(B)$ determined by the intersection points A and B of parallel lines to l_2 and l_1 through P' and the lines l_1 and l_2. This process is the same as described for the two dimensional case. To find the z coordinate we find the intersection C of a plane parallel to the xy plane and through P and l_3 and determine $c = x(C)$. We now define $[P] = (x(A), y(B), z(C)) = (a, b, c)$ as the coordinates of P given the coordinate axis x, y, z.

We define $\hat{\boldsymbol{k}} = \overrightarrow{OK}$ and $\boldsymbol{p} = x(A)\hat{\boldsymbol{i}} + y(B)\hat{\boldsymbol{j}} + z(C)\hat{\boldsymbol{k}}$ and we can represent the vector $\boldsymbol{p}$ also by the same coordinates as the point P:

$$\boldsymbol{p} = (x(A), y(B), z(C)) = (a, b, c)$$

To calculate the magnitude of $\boldsymbol{p}$ we choose $\overrightarrow{OP}$ and calculate the distance of OP. We form the triangle $OP'P$ and use the cosine rule:

$$\|\boldsymbol{p}\|^2 = OP^2 = OP'^2 + P'P^2 + 2OP' \cdot P'P \cos\theta$$

where θ is the acute angle between the line through OP' and l_3. We have:

$$\begin{aligned} OP'^2 &= a^2 + b^2 + 2ab\cos\phi \\ P'P &= c \end{aligned}$$

where ϕ it the angle between l_1 and l_2. This gives:

$$\|\boldsymbol{p}|\|^2 = OP^2 = a^2 + b^2 + c^2 + 2ab\cos\phi + 2OP'c\cos\theta \tag{1}$$

We need to derive a relation for $OP'\cos\theta$ in terms of coordinates and angles between the coordinate axes.

We proceed as follows. Draw a line through P' perpendicular to l_3 and call the intersection point P''. P'' is called the projection of P onto l_3. Then the line segment OP'' is then the projection of the line segment OP' onto l_3 and from geometry follows:

$$OP'' = OP' \cos\theta$$

We project the line segment AP' onto l_3. We have already P'' as the projection of P' and let A' be the projection of A onto l_3. AP' is parallel to l_2 by construction and we have:

$$A'P'' = AP' \cos\varphi$$

with φ the angle between l_2 and l_3

we proceed with the projection of OA onto l_3:

$$OA' = OA \cos\psi$$

where ψ is the angle between l_1 and l_3.

We have now have found how OP'' is related to the coordinates of P' and the angles between the coordinate axes.

We substitute these relations into the relation for OP'' and get:

$$OP'' = OP' \cos\theta = OA' + A'P'' = OA \cos\psi + AP' \cos\varphi = a \cos\psi + b \cos\varphi$$

We substitute this result into 1 which gives the formula for the magnitude:

$$|\boldsymbol{p}| = OP = \sqrt{a^2 + b^2 + c^2 + 2ab\cos\phi + 2ac\cos\psi + 2bc\cos\varphi}$$

If $\phi = \psi = \varphi = \frac{1}{2}\pi$, which is the case for an orthogonal Cartesian coordinate system then this formula for distance reduces to

$$|\boldsymbol{p}| = OP = \sqrt{a^2 + b^2 + c^2}$$

and for the angles:

$$\cos\alpha = \frac{a}{\sqrt{a^2 + b^2}}$$

$$\cos\gamma = \frac{c}{\sqrt{a^2 + b^2 + c^2}}$$

What we have achieved now is that we can work with coordinates (a,b,c) of a point to calculate distances and angles. From now we will denote these coordinates with symbols x, y, z.

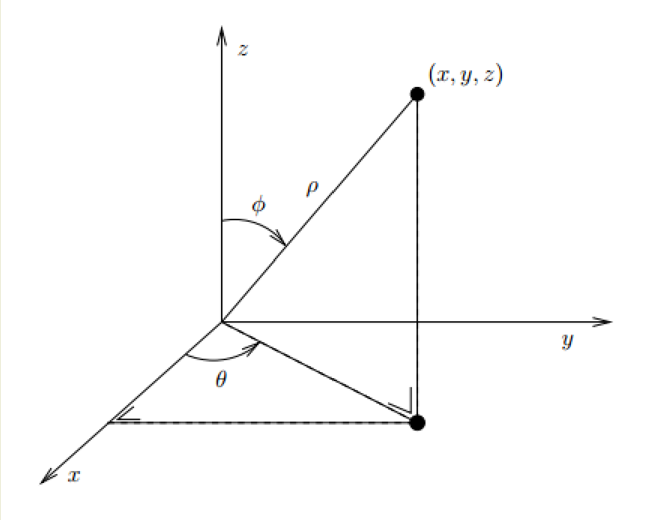

Figure 5: spherical coordinates

The following relations exist between spherical coordinates and Cartesian coordinates (note $OP' = \rho \sin \phi$):

$$
\begin{aligned}
x &= \rho \sin \phi \cos \theta \\
y &= \rho \sin \phi \sin \theta \\
z &= \rho \cos \phi
\end{aligned}
$$

In 3 dimensional space we can also utilize cylindrical coordinates:

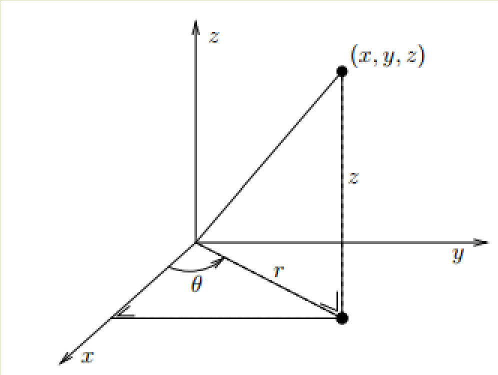

Figure 6: cylindrical coordinates

$$
\begin{aligned}
x &= r\cos\theta \\
y &= r\sin\theta \\
z &= z
\end{aligned}
$$

3.7 N Dimensional Space

For a Cartesian coordinate system we can extended the distance calculation easily to the n dimensional case $\boldsymbol{x} = (x_1, x_2, ..., x_n)$.

We assume that the distance in three dimensions is the sum of the distance in two dimensions plus the distance in the third dimension and then we see that the following pattern emerges:

$$\|\boldsymbol{x}\| = \sqrt{\sqrt{x_1^2 + x_2^2}^2 + x_3^2} = \sqrt{x_1^2 + x_2^2 + x_3^2}$$

For any arbitrary dimension n we the find:

$$\|\boldsymbol{x}\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

3.8 Vectors in physics

The directed line segment, $\overrightarrow{OP}$, is often called a *position vector*, although it is not a vector, its terminal point defines the point P, and its length and angle defines a vector. The position vector captures the concept of point and vector. In physics the position vector is often denoted as $\boldsymbol{r}$. The position vector at different times is a function of time t: $\boldsymbol{r}(t)$. The difference between two instances of time is a *displacement vector*: $\Delta\boldsymbol{r}(t)$. The instantaneous change in $\boldsymbol{r}(t)$ is the *velocity vector*:

$$\boldsymbol{v}(t) = \lim_{\Delta t \to 0} \frac{\Delta \boldsymbol{r}}{\Delta t} = \frac{d\boldsymbol{r}}{dt} = \dot{\boldsymbol{r}}(t)$$

and the instantaneous change in the velocity is the *acceleration vector*:

$$\boldsymbol{a}(t) = \lim_{\Delta t \to 0} \frac{\Delta \boldsymbol{v}}{\Delta t} = \frac{d\boldsymbol{v}}{dt} = \ddot{\boldsymbol{r}}(t)$$

Momentum, mass times velocity, is also a vector quantity, because mass is a scalar quantity and velocity a vector quantity: $\boldsymbol{p}(t) = m\boldsymbol{v}(t)$. Force is the instantaneous change of momentum and is also a vector quantity: $\boldsymbol{F}(t)$.